

**NUCLEIC ACID PROBE ASSAY METHOD AND COMPOSITION THEREFOR**

**Publication number:** JP6225799

**Publication date:** 1994-08-16

**Inventor:** INADA ASAMI; DAIMON KATSUYA; HAYASHI SATOKO

**Applicant:** TOYO BOSEKI

**Classification:**

- **International:** **C12Q1/68; C12N15/09; C12N15/09; C12Q1/68; C12N15/09; C12N15/09; (IPC1-7): C12Q1/68**

- **European:**

**Application number:** JP19930014826 19930201

**Priority number(s):** JP19930014826 19930201

[Report a data error here](#)

**Abstract of JP6225799**

**PURPOSE:**To obtain a highly sensitive detection method capable of preventing nonspecific adsorption of a labeled substance into a solid phase in a nucleic probe assay method and a composition therefor.

**CONSTITUTION:**In a method for detecting a nucleic acid to be analyzed by previously binding a capturing probe to a solid phase and capturing the nucleic acid to be analyzed with the capturing probe, the nucleic acid probe assay method is characterized by treating the solid phase to which the capturing probe is bound with a reagent containing a mononucleotide and/or a mononucleoside. In a method for binding a nucleic acid to be analyzed in a sample to a solid phase and detecting the nucleic acid to be analyzed by a labeled probe, the nucleic acid probe assay method is characterized by treating the solid phase to which the nucleic acid to be analyzed is bound with a reagent containing a mononucleotide and/or mononucleoside and then reacting a labeled probe.

.....  
Data supplied from the **esp@cenet** database - Worldwide

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A)

昭62-25799

⑬ Int.Cl.<sup>4</sup>

識別記号

庁内整理番号

⑭ 公開 昭和62年(1987)2月3日

G 10 L 7/08

A-8221-5D

審査請求 未請求 発明の数 1 (全7頁)

⑮ 発明の名称 音声認識装置

⑯ 特 願 昭60-166191

⑰ 出 願 昭60(1985)7月27日

⑱ 発 明 者 納 田 重 利 東京都品川区北品川6丁目7番35号 ソニー株式会社内

⑲ 出 願 人 ソ ニ ー 株 式 会 社 東京都品川区北品川6丁目7番35号

⑳ 代 理 人 弁理士 杉浦 正知

# 明 細 書

## 〔産業上の利用分野〕

### 1. 発明の名称

音声認識装置

この発明は、例えば話者の音声単語単位で認識するのに適用される音声認識装置に関する。

### 2. 特許請求の範囲

入力音声信号が複数チャンネルの周波数スペクトルに変換され、上記複数チャンネルの周波数スペクトルの時系列データが入力される音声認識装置において、

上記時系列データの各フレームのスペクトルデータに関して所定のチャンネルより低い全てのチャンネルの上記スペクトルデータの第1の平均値を算出すると共に、上記所定のチャンネルより高い全てのチャンネルの上記スペクトルデータの第2の平均値を算出し、上記第1の平均値と上記第2の平均値との平均値を上記所定のチャンネルにおける傾向値として算出し、上記傾向値又は適当なオフセットが付加された上記傾向値を基準レベルとして二値化処理を行うことを特徴とする音声認識装置。

### 3. 発明の詳細な説明

## 〔発明の概要〕

この発明は、音声認識装置において、種々の原因により変動するスペクトルの傾向を補正するための傾向値を算出し、この傾向値に基づいてスペクトルの傾向を平坦化することにより話者の個人差や周囲のノイズ等に影響されることがないようにして認識率の向上を図ると共に、二値化処理を行うことによりメモリ容量を低減させ、計算処理時間を短縮するようにしたものである。

## 〔従来の技術〕

従来の音声認識装置としては、例えば音声入力部としてのマイクロホン、前処理回路、音響分析器、特徴データ抽出器、登録パターンメモリ及びパターンマッチング判定器等により構成されるものが知られている。

この音声認識装置は、マイクロホンから入力される音声信号を前処理回路において、音声認識に必要とされる帯域に制限し、A/D変換器によりデジタル音声信号とし、このデジタル音声信号を音響分析器に供給する。

そして、音響分析器において、音声信号を周波数スペクトルに変換し、例えば対数軸上で一定間隔となるようにN個の周波数を代表値とする周波数スペクトルを算出、フレーム間隔毎にNチャンネルのスペクトルデータにより構成されるフレームデータの特徴データ抽出器に供給する。特徴データ抽出器は、隔り合うフレームデータの距離を計算し、夫々のフレーム間距離の総和により、音声信号の始端フレームから終端フレームまでのN次元ベクトルの軌跡長を求める。最も語数が多く長い音声の場合に特徴を抽出するのに必要な所定の分割数をもって軌跡長を等分割し、その分割点に対応したフレームデータのみを特徴データとして抽出する。

この特徴データを登録時においては、登録パ

ターンメモリに供給して登録特徴データブロック(標準パターン)として記憶し、認識時においては、入力音声信号を前述した処理により、入力特徴データブロック(入力パターン)とし、パターンマッチング判定器に供給する。そしてパターンマッチング判定器において、入力特徴データブロックと登録特徴データブロックとの間でパターンマッチングを行う。

パターンマッチング判定器は、登録特徴データブロックを構成するフレームデータと入力特徴データブロックを構成するフレームデータとの間でフレーム間距離を計算し、フレーム間距離の総和をマッチング距離とする。他の登録特徴データブロックに關しても同様にマッチング距離を算出して、マッチング距離が最小で十分に距離に近いものと判断される登録特徴データブロックに対応する単語を認識結果として出力する。

(発明が解決しようとする問題点)

しかし、音声信号の周波数スペクトルは話者の

個人差及び周囲ノイズ等の差人によってその傾向が大きく変化するもので、この傾向を正規化しないと認識率が極めて低下する。

例えば第6図Aに示すフレームデータが第6図B示すようなスペクトル傾向を持つノイズにより変形され、第6図Cに示すようなフレームデータとされたとする。パターンマッチング判定器において、第6図Aに示すフレームと第6図Cに示すフレームとの距離が求められると、そのフレーム間距離は大きな値として計算され、マッチング距離が大きな誤差を含んだものとされて誤認識される可能性が高くなる。このため、スペクトルの傾向変動を補正して、話者の個人差や周囲ノイズに影響されることがないようにスペクトルの傾向を平坦化(正規化)することが提案されている。

例えば最小二乗法等でスペクトル傾向を一次関数で推定し正規化する手法や所定のチャンネル幅間で部分的に平均化した補正関数で正規化する手法が提案されている。しかしながら、前者の場合は、計算が複雑なばかりか傾向が曲線を描く場合

に適用することができず、また、後者の場合は、スペクトルエンベロップがなめらかな場合に適用することができない欠点を有するものであった。

従って、この発明の目的は、簡単にかつ高速に任意のスペクトル傾向を正確に正規化することができる手段を有した音声認識装置を提供することにある。

また、従来の音声認識装置においては、音響分析器から出力されるフレームデータが特徴データ抽出器を介してそのまま登録特徴データブロックとして登録パターンメモリに記憶されるため、登録パターンメモリのメモリ量が膨大なものとなる問題点があった。これと共に、パターンマッチング時においても、データ量に応じてその計算処理時間が長くなる問題点があった。

従って、この発明の他の目的は、フレームデータを二値化することにより、登録パターンメモリの容量を低減でき、また、マッチング処理時間の短縮を図ることができる音声認識装置を提供することにある。

〔問題点を解決するための手段〕

この発明は、複数チャンネルの周波数スペクトルに変換され、複数チャンネルの周波数スペクトルの時系列データが入力される音声認識装置において、

時系列データの各フレームのスペクトルデータに関して所定のチャンネルより低い全てのチャンネルのスペクトルデータの第1の平均値を算出すると共に、所定のチャンネルより高い全てのチャンネルのスペクトルデータの第2の平均値を算出し、第1の平均値と第2の平均値との平均値を所定チャンネルにおける傾向値として算出し、傾向値又は適当なオフセットが付加された傾向値を基準レベルとして二値化処理を行うことを特徴とする音声認識装置である。

〔作用〕

スペクトルの傾向を正規化する手段としてスペクトル傾向正規化器6が設けられると共に、二値

化回路8が設けられ、スペクトル傾向正規化器6において、時系列フレームデータのフレーム毎に、チャンネル1から所定のチャンネル $n$  ( $1 \leq n \leq N$ ) までのスペクトルデータの平均値が求められると共に、所定のチャンネル $n$ から最大チャンネル $N$ までのスペクトルデータの平均値が求められる、夫々の平均値の更に平均値が求められて所定のチャンネル $n$ に関する傾向値とされ、各チャンネルのスペクトルデータと対応する傾向値との間において夫々減算処理がなされることにより固有の特徴的なスペクトルデータが保存されながら、スペクトル傾向が平坦化される。二値化回路8において、スペクトルデータが二値データとされ、この二値データが登録パターンメモリ10及びパターンマッチング判定器11に供給され、この二値データに基づいてパターンマッチングが行われる。

〔実施例〕

以下、この発明の一実施例を図面を参照して説明する。第1図は、この発明の一実施例を示すも

ので、第1図において1が音声入力部としてのマイクロホンを示している。

マイクロホン1からのアナログ音声信号がフィルタ2に供給される。フィルタ2は、例えばカットオフ周波数7.5 KHzのローパスフィルタであり、音声信号がフィルタ2において、音声認識に必要とされる7.5 KHz以下の帯域に制限され、この音声信号がアンプ3を介してA/D変換器4に供給される。

A/D変換器4は、例えば、サンプリング周波数12.5 KHzで動作するA/D変換器である。音声信号がA/D変換器4において、アナログデジタル変換されて、8ビットのデジタル信号とされ、音響分析器5に供給される。

音響分析器5は、音声信号を周波数スペクトルに変換して、例えば $N$ チャンネルのスペクトルデータ列を発生するものである。音響分析器5において、音声信号が演算処理により周波数スペクトルに変換され、例えば、対数軸上で一定間隔となる $N$ 個の周波数を代表値とするスペクトルデータ

列が得られる。従って、音声信号が $N$ チャンネルの離散的な周波数スペクトルの大きさによって表現される。そして、単位時間（フレーム周期）毎に $N$ チャンネルのスペクトルデータ列が一つのフレームデータとして出力される。即ち、フレーム周期毎に音声信号が $N$ 次元ベクトルにより表現されるパラメータとして切り出され、スペクトル傾向正規化器6に供給される。

例えば、音声区間の終端に対応するフレームを1とした場合、第2図に示すように、各々がチャンネル1～チャンネル $N$ のデータにより構成されるフレームデータが1フレームから1フレームまでスペクトル傾向正規化器6に供給される。

スペクトル傾向正規化器6は、傾向値計算回路12及び減算器13により構成されている。このスペクトル傾向正規化器6において、順次供給されるフレームデータ毎にスペクトルデータの傾向正規化処理がなされる。

傾向値計算回路12において、フレームデータを構成する各チャンネルのスペクトルデータに關

して傾向変動を補正する傾向値F、が下記の式により算出される。

$$F = \frac{(N+1-n) \sum_{i=1}^n S_i + n \sum_{i=n+1}^N S_i}{2(N+1-n) \cdot n}$$

つまり、チャンネル1から所定のチャンネルn(1 ≤ n ≤ N)までのスペクトルデータの平均値が求められると共に、所定のチャンネルnから最大チャンネルNまでのスペクトルデータの平均値が求められる。更に夫々の平均値の平均値が求められ、この平均値が傾向値Fとされる。N個の傾向値データが演算器13に供給される。

演算器13において、各チャンネルのスペクトルデータと対応する傾向値データとが演算される。この演算処理によりスペクトル傾向が平坦化され、話者の個人差及び周囲ノイズ等に影響されることがないようにスペクトル傾向が正規化される。1フレームから1フレームまで全てのフレームに関して同様に傾向正規化処理がなされ、傾向正規化されたフレームデータが特徴データ抽出器7に供

給される。

特徴データ抽出器7において、隣り合うフレームデータの距離が計算される。例えば、各チャンネルに関してスペクトルデータの差の絶対値が夫々求められ、その総和がフレーム間距離とされる。

更に、フレーム間距離の総和が求められ、音声信号の始端フレームから終端フレームまでのN次元ベクトルの軌跡長が求められる。そして最も語数が多く長い音声の場合に特徴を抽出するのに必要な所定の分割数をもって軌跡長が等分割される。分割点の夫々に対応したフレームデータのみが特徴データとして抽出され、話者の音声の発生速度変動に影響されることがないように時間軸が正規化される。

特徴データ抽出器7により抽出されたフレームデータが二値化回路8に供給される。二値化回路8は、入力端子14を有しており、入力端子14を介して適当に設定された基準レベルデータが供給される。この基準レベルデータとフレームデータを構成する夫のスペクトルデータとの比較がな

され、基準レベルデータより大きな値のスペクトルデータが「1」とされ、基準レベルデータより小さな値のスペクトルデータが「0」とされて二値化される。この二値データがモード切替回路9に供給される。

この二値データが登録時においては、モード切替回路9を介して登録パターンメモリ10に供給され、登録特徴データブロックとして記憶される。認識時においては、入力音声信号が前述した処理により二値データ(入力特徴データブロック)とされ、この二値データがパターンマッチング判定器11に供給される。入力特徴データブロックと全ての登録特徴データブロックとの間において、パターンマッチングが行われる。

即ち、パターンマッチング判定器11において、登録パターンメモリ10から順次供給される登録特徴データブロックを構成するフレームと入力特徴データブロックを構成するフレームとの間において、フレーム間距離が求められ、その総和がマッチング距離とされる。そして全ての登録特徴デ

ータブロックに関して求められたマッチング距離のうちで最小でかつ十分に距離が近いものと判断される登録特徴データブロックに対応する単語が認識結果とされる。

上述のこの発明の一実施例におけるスペクトル傾向正規化器6の動作を第3図に示すフローチャートを参照して説明する。

音響分析器5から順次フレームデータがスペクトル傾向正規化器6に供給され、各フレーム毎にステップ①～⑩の処理が行われる。

まず、チャンネル番号を示す変数nが1に初期設定される(ステップ①)。ステップ②において、チャンネル1に関する補正関数の計算処理がなされ、傾向変動を補正する傾向値F<sub>1</sub>が

$$F_1 = \frac{(N+1-1) \sum_{i=1}^1 S_i + 1 \cdot \sum_{i=1}^N S_i}{2(N+1-1) \cdot 1}$$

により求められる。

そして、正規化処理がステップ③においてなされ、チャンネル1のスペクトルデータS<sub>1</sub>から傾

向値 $F_1$ が減算され、この減算結果がチャンネル1のスペクトルデータ $S_1$ とされる。

ステップ⑩において、チャンネル番号を示す変数 $n$ と最大チャンネル数 $N$ との比較がなされ、 $n$ がインクリメントされて( $n=2$ ) (ステップ⑨)とされ、チャンネル2に関する計算処理に移行する。

傾向変動を補正する傾向値 $F_2$ が

$$F_2 = \frac{(N+1-2) \cdot \sum_{i=1}^N S_i + 2 \cdot \sum_{i=2}^N S_i}{2(N+1-2) \cdot 2}$$

により求められ(ステップ⑨)、チャンネル2のスペクトルデータ $S_2$ から傾向値 $F_2$ が減算され、この減算結果がチャンネル2のスペクトルデータ $S_2$ とされる。(ステップ⑩)。

更に $n$ がインクリメントされながら、上述したステップ⑨～⑩の処理が繰り返行われ、所定チャンネルに関しての傾向値 $F_n$ が

$$F_n = \frac{(N+1-n) \cdot \sum_{i=1}^N S_i + n \cdot \sum_{i=n}^N S_i}{2(N+1-n) \cdot n}$$

されながらスペクトル傾向が平坦化される。

また、第4図Cに示す傾向正規化処理がなされたフレームデータが特徴データ抽出器7を介して二値化回路8に供給されると、基準レベルにより二値化され、(0.1,1,1,0,1,1,0,0,0,1,0,0,0,0)の二値データとされる。

第5図は、この発明の他の実施例を示し、二値化回路8において、スペクトルの傾向正規化処理と二値化処理を同時に行う構成とされ、上述の実施例と対応する部分には、同一の符号が付されている。音響分析器5から順次フレームデータが傾向値計算回路12に供給される。

傾向値計算回路12において、フレームデータを構成する各チャンネルのスペクトルデータに関する傾向変動を補正する傾向値 $F$ が算出され、傾向値データが加算器15に供給される。加算器15には、入力端子16から適当に設定されたオフセットデータが供給される。加算器15において、傾向値データとオフセットデータとの加算処理がなされ、この加算結果が基準レベルデータと

により求められ(ステップ⑨)、所定チャンネルのスペクトルデータ $S$ から傾向値 $F$ が減算され、この減算結果が所定チャンネルのスペクトルデータ $S$ とされる。チャンネル番号を示す変数 $n$ が最大チャンネル数 $N$ とされ、最大チャンネルに関して傾向正規化処理がなされると、一つのフレームに関する計算が終了する。

例えば、第4図Aに示すようなチャンネル1～チャンネル16の16個のスペクトルデータにより構成され、各チャンネルのスペクトルデータの大きさが(8, 12, 16, 17, 12, 14, 18, 16, 12, 10, 6, 12, 9, 8, 6, 5)とされるフレームデータについて説明する。このフレームデータの場合には、上述の処理によりチャンネル1～チャンネル16までの傾向値 $F$ は第4図Bに示すものとなる。この傾向値 $F$ を基準としてスペクトルデータが平坦化され、傾向正規化されたスペクトルデータは、第4図Cに示すものとなる。このように全てのフレームに関して傾向正規化処理がなされ、固有の特徴的なスペクトルデータが保存

して二値化回路8に供給される。

二値化回路8において、基準レベルデータとフレームデータを構成する各のスペクトルデータとの比較がなされ、基準レベルデータより大きな値のスペクトルデータが「1」とされ、基準レベルデータより小さな値のスペクトルデータが「0」とされて二値化され、この二値データがモード切替回路9に供給される。

この二値データが登録時においては、モード切替回路9を介して登録パターンメモリ10に供給され、登録データブロックとして記憶される。認識時においては、入力音声信号が前述した処理を経ることにより二値データとされ、この二値データがパターンマッチング判定器17に供給され、入力データブロックとされる。

パターンマッチング判定器17において、話者の音声の発生速度変動によるデータブロックの構成フレーム数の増減を吸収する例えばDPマッチングにより、登録データブロックと入力データブロックとのマッチング距離が算出される。そして、

全ての登録データブロックに関して求められたマッチング距離のうちで最小でかつ十分距離が近いものと判断される登録データブロックに対応する単語が認識結果とされる。

尚、この発明の他の実施例においては、傾向値データにオフセットデータを加算する構成について説明したが、スペクトルデータからオフセットデータを減算する構成とし、傾向値データを基準レベルデータとして二値化回路に供給して二値化処理を行うようにしても良い。また、この発明は、ハードワイヤードの構成に限らず、マイクロコンピュータ又はマイクロプログラム方式を用いてソフトウェアにより処理を行うようにしても良い。

#### (発明の効果)

この発明では、スペクトルの傾向を正規化することにより、固有の特徴的なスペクトルデータが保存されながら、スペクトル傾向が平坦化される。また、この発明では、二値化回路において、スペクトルデータが二値データとされ、この二値データが登録パターンメモリ及びパターンマッチング

判定器に供給され、この二値データに基づいてパターンマッチングが行われる。

従って、この発明に依れば、簡単でかつ高速に任意のスペクトル傾向を正確に正規化することができ、計算処理時間が短縮されると共に認識率が向上される。

また、この発明に依れば、フレームデータのスペクトルの傾向が正規化されているためフレームデータを正確に二値化することができ、例えば1個のスペクトルデータが8ビットで表される場合には、登録パターンメモリの容量を1/8に低減できると共に、マッチング処理時間が大幅に短縮される。

#### 4. 図面の簡単な説明

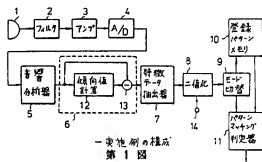
第1図はこの発明の一実施例の構成のブロック図、第2図はこの発明の一実施例における時系列フレームデータのデータ構成の説明に用いる略線図、第3図はこの発明の一実施例におけるスペクトル傾向正規化器の動作説明に用いるフローチャ

ート、第4図A、第4図B及び第4図Cはこの発明の一実施例におけるスペクトル傾向正規化器の動作説明に用いる略線図、第5図はこの発明の他の実施例の構成のブロック図、第6図は従来の音声認識装置の説明に用いる略線図である。

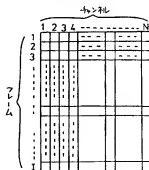
#### 図面における主要な符号の説明

- 1 : マイクロホン、 5 : 音響分析器、
- 6 : スペクトル傾向正規化器、
- 7 : 特徴データ抽出器、
- 8 : 二値化回路、 9 : モード切替回路
- 10 : 登録パターンメモリ、
- 11、17 : パターンマッチング判定器。

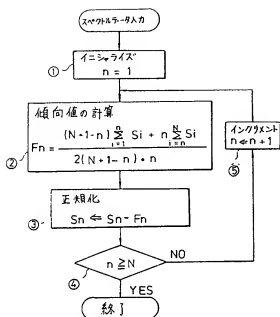
代理人 弁理士 杉 浦 正 知



一実施例の構成  
第1図

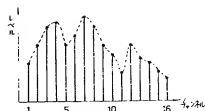


時系列フレームデータ  
第2図

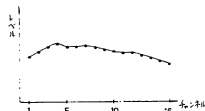


動作説明用のフローチャート

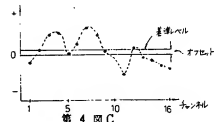
第3図



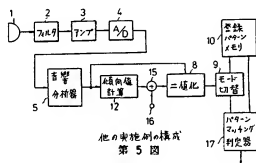
第4図A



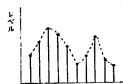
第4図B



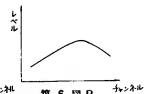
第4図C



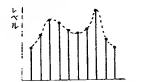
第5図



第6図A



第6図B



第6図C